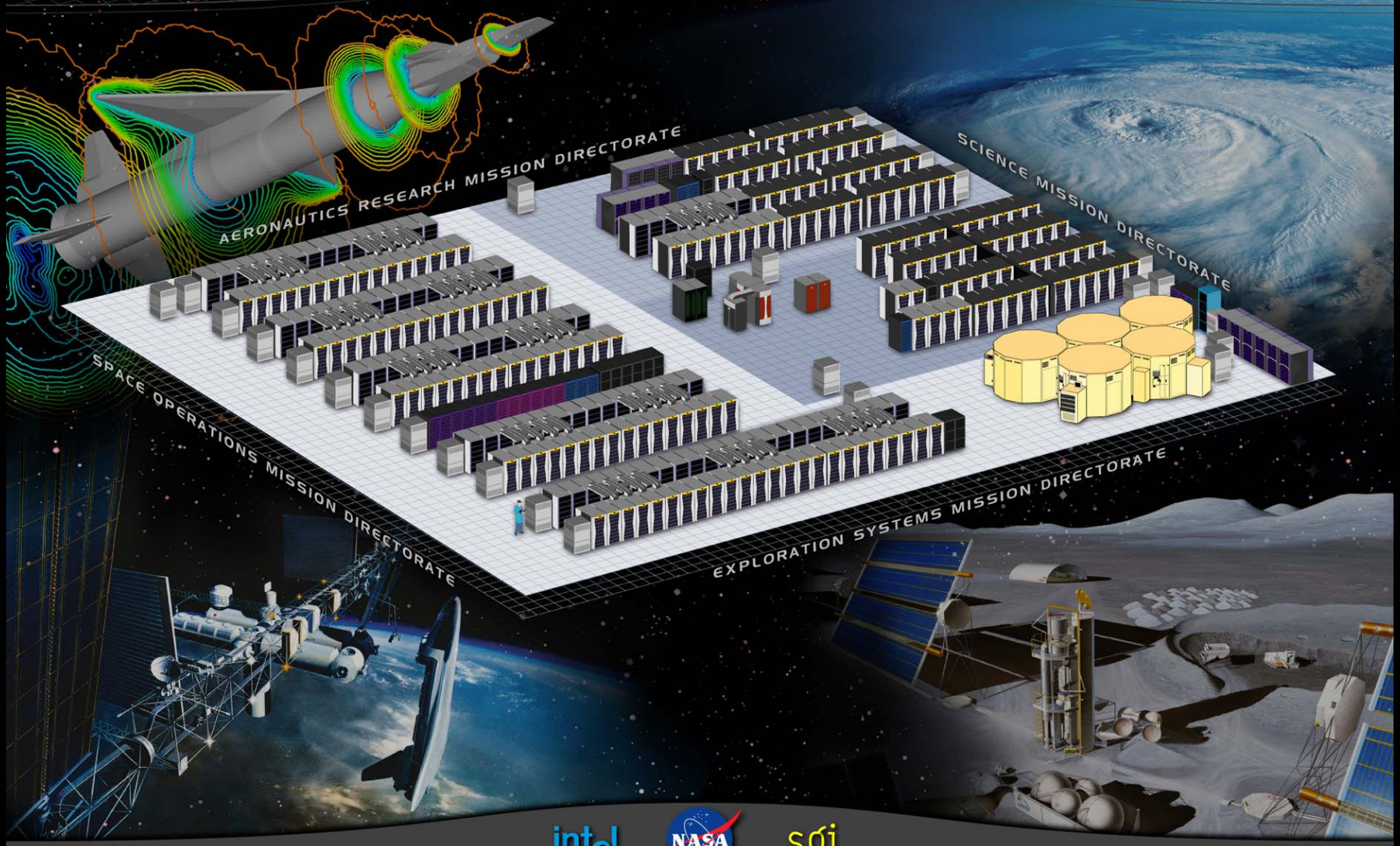
PROJECT COLUMBIA

SUPPORTING NASA'S MISSION DIRECTORATES



intel. NASA sgi

Project Columbia
From 0 to 60 Teraflops in 120 Days

Fastest production supercomputer in the world!

Project Management Preview

- Many NASA projects face challenging goals, firm deadlines, tight schedules, uncertain support, and major technical hurdles
- The good news: this is not an *automatic* recipe for project failure
- Project Columbia overcame these and other challenges to achieve dramatic successes
- This talk discusses merging the right amount of management controls and techniques with the right people to achieve a positive outcome

Agenda

- Project Columbia Overview
- Tools and Techniques
- Project Execution
- Results
- Summary
- Q&A

Project Columbia Overview

- NASA-industry partnership, formed May 2004, to significantly enhance NASA and national supercomputing
- Seized opportunity to meet specific and compelling industry and NASA objectives
- Matches NASA needs and expertise, and was not funded by other supercomputing pathfinding programs
- Leadership system built from proven technologies
- Provided NASA with immediate and increasing supercomputing power
- Supports inter-agency supercomputing task force findings
- In < 1 year, Project Columbia wildly successful for NASA
- Biggest impact is yet to come!

COLUMBIA PROJECT AT NAS


Floor Preparation


Repositioning Old Equipment


Meetings and More Meetings


New Network Racks


Four of the 10,240 Processors


Ready for Science


New Power Panels


New Cooling Piping


440 TB of RAID Storage


Cabling a New Arrival


The View From Above

Internal Networks
Internode Comm - Infiniband
Data Transfer - Gigabit Ethernet
Hi-Speed - 10 Gigabit Ethernet

Systems - SGI Altix 3700 and 3700-BX2
Processors - 10.240 Intel Itanium 2
Global Shared Memory - 20 Terabytes

Front-End - SGI Altix 3700 (64 proc)
Online Storage - 440 Terabytes RAID
Offline Storage - 6 Petabyte STK Silos

Interactive Visual Supercomputing

Application-specific exploration and analysis of complex, large datasets

Tools and Techniques of Project Success

- Team Formation
- Planning
- Communication
  - Scheduling / Project Management Tool
  - Regular Meetings
  - Tracking
- Constant Improvement
  - Vendor Processes
  - Internal Processes
- Customer Focus

Team Formation

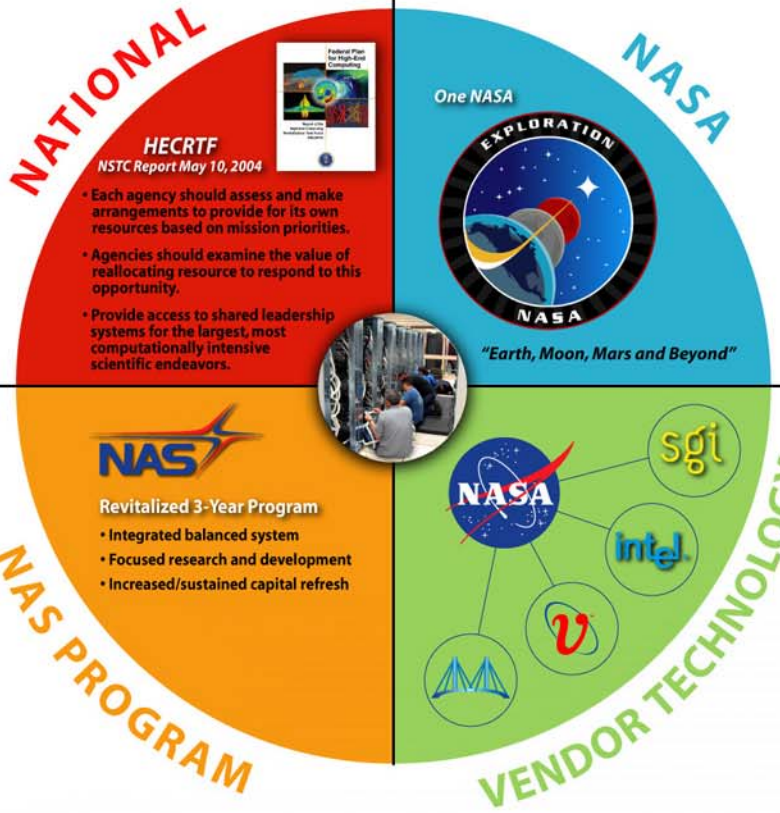
- Inter-organizational team
  - NAS, Intel, SGI, Voltaire, Melanox
  - ARC support organizations
- Focused internal team
  - Diverse opinionated experts
  - Focus on areas of expertise
  - Daily interactions
  - The right amount of methodology
  - Tight tracking of issues
  - Encouragement of all ideas
  - Clear goals
  - Continuation of build
  - No fault

NASA ADVANCED SUPERCOMPUTING MISSION

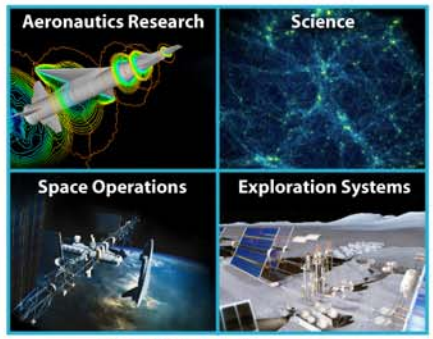
An integrated modeling and simulation environment enabling NASA and its industry and academic partners to accelerate design cycle time, conduct extensive parameter studies of multiple mission scenarios, and increase safety during the entire life cycle of exploration missions, while satisfying the tight time constraints of fast-paced NASA exploration system design and acquisition cycles.

"LEADERSHIP" SYSTEMS

Success through Partnerships



MISSION CRITICAL APPLICATIONS

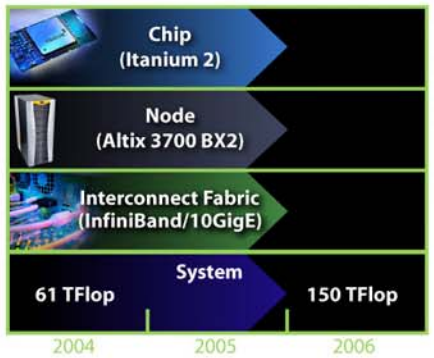


NASA's Mission Directorates

INTEGRATED ENVIRONMENTS

Next Generation Codes & Algorithms	Rapid Deployment
OVERFLOW - *STS 107* Honorable Mention, 1998 INS3D - *Pump Analysis* NASA Software of the Year, 1994 CART3D - *STS 107* NASA Software of the Year, 2002	- Systems - Networks - Archives
Data Analysis & Visualization	**Computer Science Tools & Expertise**
- Hyperwall visualization systems - Process large data sets - Interactive real-time visualization	- Program development tools - Automatic multilevel parallelization - Scaling, porting, and optimization

SUPERCOMPUTING



Columbia System

Planning

- Initial discussions between Division Chief and Intel
- System planning discussions with Intel, NASA and various candidate vendors
- System planning with Intel, NASA, and SGI
- Selling the Program to NASA
- OMB and the hill
  - Exibit300
  - Continuing discussions with SGI
- The detailed execution plan

Communication

- Scheduling / Project Management Tool
  - MicroPlanner Manager/X-Pert
  - Pert and Gantt Charts
- Controlled Diagrams
  - Floor Layout
  - Delivery Schedule
  - Network and Cabinet diagrams
- Standing Meetings
  - Daily Standups
  - Weekly Status
  - Coordination meetings with customers and vendor partners
- Tracking
  - Schedule and layout updates

Configuration Management
Columbia Configuration Chart

Columbia Configuration Chart		Columbia-1	Columbia-2	Columbia-3	Columbia-4	Columbia-5	Columbia-6	Columbia-7	Columbia-8	Columbia-9	Columbia-10	Columbia-11	Columbia-12	Columbia-13	Columbia-14	Columbia-15	Columbia-16	Columbia-17	Columbia-18	Columbia-19	Columbia-20	Columbia (front end)	= Production System = Application Devel = Systems Devel = Hardw Install / Test	Node Names & Color Key
Hardware Configuration	CPU Speed (gHz)	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.6	1.5	1.5	1.5	1.5	1.5	1.5	CPU Speed (gHz)	Hardware Configuration
	Cache Size (mB/p)	6	6	6	6	6	6	6	6	6	6	6	6	6	6	9	6	6	6	6	6	6	Cache Size (mB/p)	
	Prom Level	3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.66	3.66	3.66	3.66	3.50	Prom Level	
	CPU Configuration	512	512	512	512	512	512	512	512	512	512	512	512	512	512	512	512	512	512	512	512	64	CPU Configuration	
	Mem Module Size (gB)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		Mem Module Size (gB)	
	Total Mem Size (tB)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.12	Total Mem Size (tB)	
	FiberCh 2-port adapters	4	4	4	4	4	4	4	4	4	4	3	4	4	4	4	4	4	4	4	4	1	FiberCh 2-port adapters	
	gigE 2-port adapters	0	0	5	0	0	0	0	0	0	0	0	0	5	5	0	5	4	4	5	5	1	gigE 2-port adapters	
	10gigE adapters	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	10gigE adapters	
	Infiniband adapters	4	4	4	4	4	4	4	4	4	4	4	4	8	8	8	8	8	8	8	8	0	Infiniband adapters	
	IX-brick	4	4	2	4	4	4	4	4	3	3	4	4	4	4	4	4	4	4	3	4	1	IX-brick	
Software Configuration	Linux OS Level	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	2.4.21	Linux OS Level	Software Configuration
	Propack Level	3.2c21	3.2c21	3.2c21	3.2c21	3.2c21	3.2c21	3.2c21	3.2c21	3.2c21	3.2c21	3.2c21	3.2c21	3.2c21	3.2c21	3.3r2	3.2c21	3.2c21	3.2c21+	3.2c21+	3.2c21+	3.2c21+	Propack Level	
	PBS Level	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	5.3.2	PBS Level	
	MPT level	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	1.9.1.0	MPT level	
	Default Intel Compiler	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	7.1.032	Default Intel Compiler	
	SCSL	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0.0	SCSL	
	CXFS Version																						CXFS Version	
	SSH Version	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	3.9p1	SSH Version	
	Pending Patch																						Pending Patch	
History & Status	Model	3700	3700	3700	3700	3700	3700	3700	3700	3700	3700	3700	3700	BX2	BX2	BX2	BX2	BX2	BX2	BX2	BX2	3700	Model	History & Status
	Ship Date	Jun 25	Jun 25	Jul 30	2003	Aug 17	Aug 13	Sep 24	Sep 15	Sep 21	Sep 17	Aug 30	Sep 24	Sep 24	Sep 24	Sep 17	Sep 10	Oct 12	Oct 8	Oct 8	Oct 1		Ship Date	
	Serial Number	N0000772	N0000771	N0000793	N0000384	N0000812	N0000809	N0000833	N0000814	N0000832	N0000821	N0000816	N0000848	N0000830	N0000850	N0000817	N0000823	N0000855	N0000854	N0000849	N0000853	N0000726	Serial Number	
	Install Date	Jun 30	Jun 28	Aug 2	2003	Aug 20	Aug 16	Sep 24	Sep 1	Sep 23	Sep 17	Sep 10	Sep 24	Sep 27	Sep 27	Sep 20	Sep 13	Oct 12	Oct 12	Sep 27	Sep 27		Install Date	
	Floor Location	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	HSP-B	Floor Location	
	PDU connection	3,23,31	3,23,31	1,2	1,2	4,5	4,5	6,7	6,7	8,9	8,9	10	22,24	12	13	14	15	16	17	18	19		PDU connection	
	System "Name"																						System "Name"	
	Node Usage	SpOps	Science	AppDv	SysDv	Explor	Aero	SpOps	Science	SysDv	Sys Dv	Explor	Aero	AppDv	AppDv	SysDv	AppDv	2048 Dv	2048 Dv	2048 Dv	2048 Dv	FtEnd	Node Usage	

Version 13 - 01/27/05
DGR

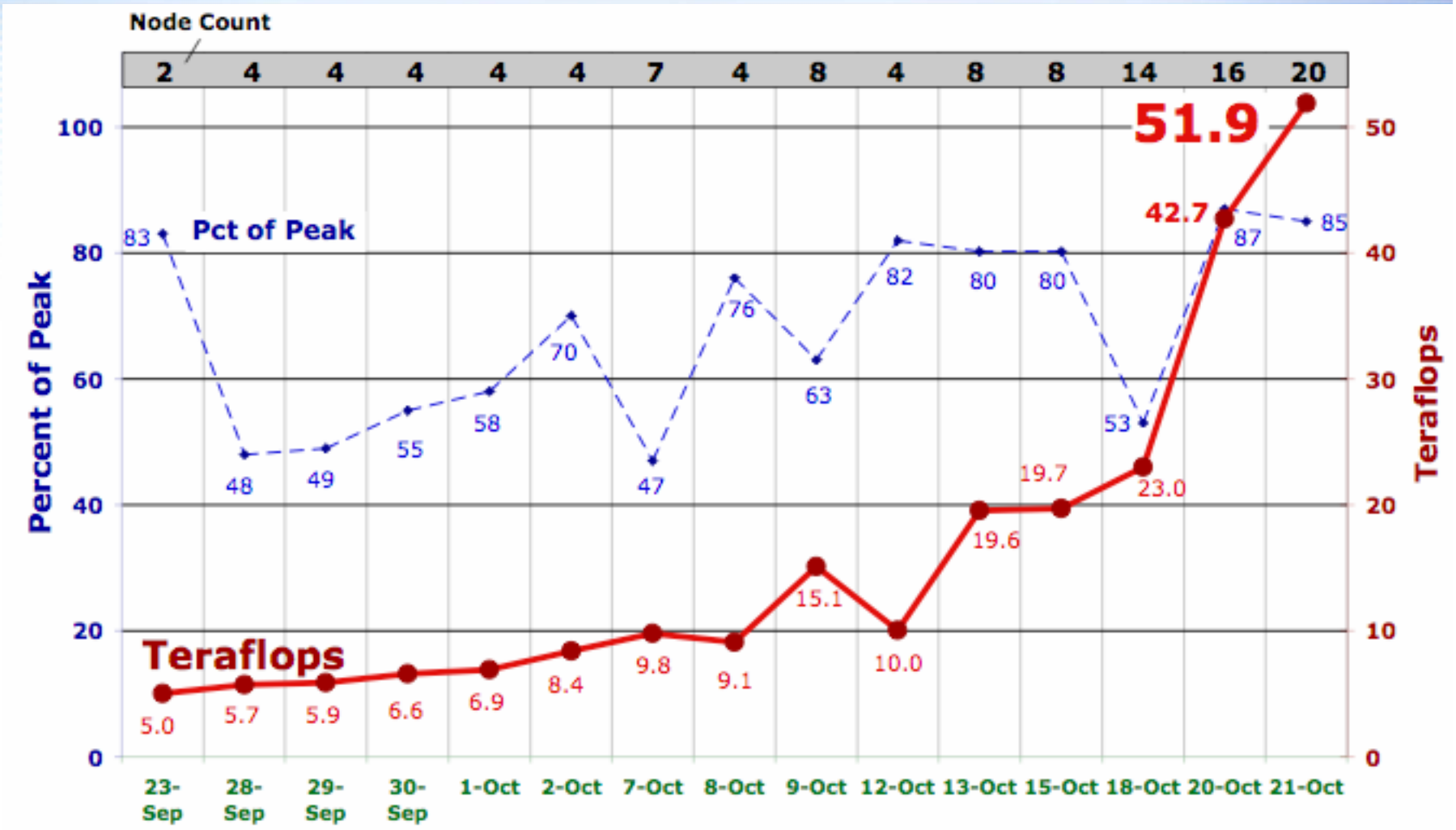
Version 13 - 01/27/05

Chart Detail

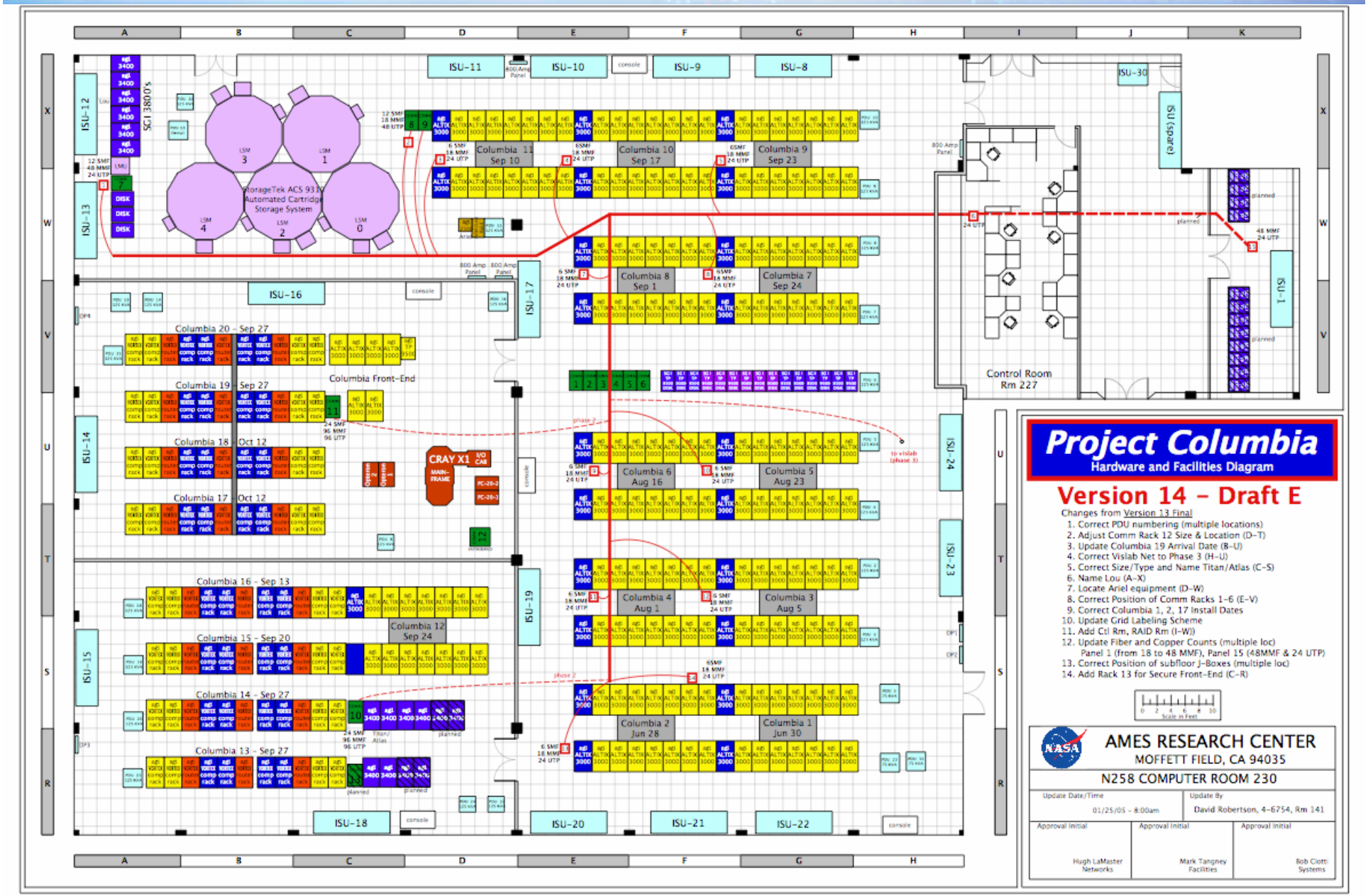
- Systems were color coded based on utilization
- Rows depicted controlled hardware and software as well as status
- Provided easy access to entire staff depicting the current systems configuration
- All changes to this were accomplished through the CCB

Columbia Configuration Chart		Columbia-1	Columbia-2	Columbia-3	Columbia-4
Hardware Configuration	CPU Speed (gHz)	1.5	1.5	1.5	1.5
	Cache Size (mB/p)	6	6	6	6
	Prom Level	3.50	3.50	3.50	3.5
	CPU Configuration	512	512	512	51
	Mem Module Size (gB)	1	1	1	1
	Total Mem Size (tB)	1	1	1	1
	FiberCh 2-port adapters	4	4	4	4
	gigE 2-port adapters	0	0	5	0
	10gigE adapters	1	1	1	1
	Infiniband adapters	4	4	4	4
	IX-brick	4	4	2	4
Software Configuration	Linux OS Level	2.4.21	2.4.21	2.4.21	2.4.
	Propack Level	3.2c21	3.2c21	3.2c21	3.2c
	PBS Level	5.3.2	5.3.2	5.3.2	5.3
	MPT level	1.9.1.0	1.9.1.0	1.9.1.0	1.9.
	Default Intel Compiler	7.1.032	7.1.032	7.1.032	7.1.0
	SCSL	1.5.0.0	1.5.0.0	1.5.0.0	1.5.0
	CXFS Version				
	SSH Version	3.9p1	3.9p1	3.9p1	3.9p
	Pending Patch				
History & Status	Model	3700	3700	3700	370
	Ship Date	Jun 25	Jun 25	Jul 30	200
	Serial Number	N0000772	N0000771	N0000793	N0000
	Install Date	Jun 30	Jun 28	Aug 2	200
	Floor Location	1	2	3	4
	PDU connection	3,23,31	3,23,31	1,2	1,
	System "Name"				
	Node Usage	SpOps	Science	AppDv	Sys

Linpack Results



Main Computer Floor



Enabled Changes to occur

- Each version was approved by networks, facilities and systems with project concurrence
- The drafts carried the changes since the last approved version.
- Like the configuration chart, this was an effective means of communicating and verifying the current computer room state
- Drawing was to scale and was used by the team for the placement of all of the equipment

Constant Improvement

- Build / Evaluate / Improve / Build
  - SGI System Delivery
  - AMTI / CSC Systems Team
  - Facility Group
  - Networks
- Eye on the customer
  - Going the extra step to prevent or minimize customer impact
  - Regular customer interactions
  - Teams formed centered on customer requirements
  - Communicate, communicate, communicate

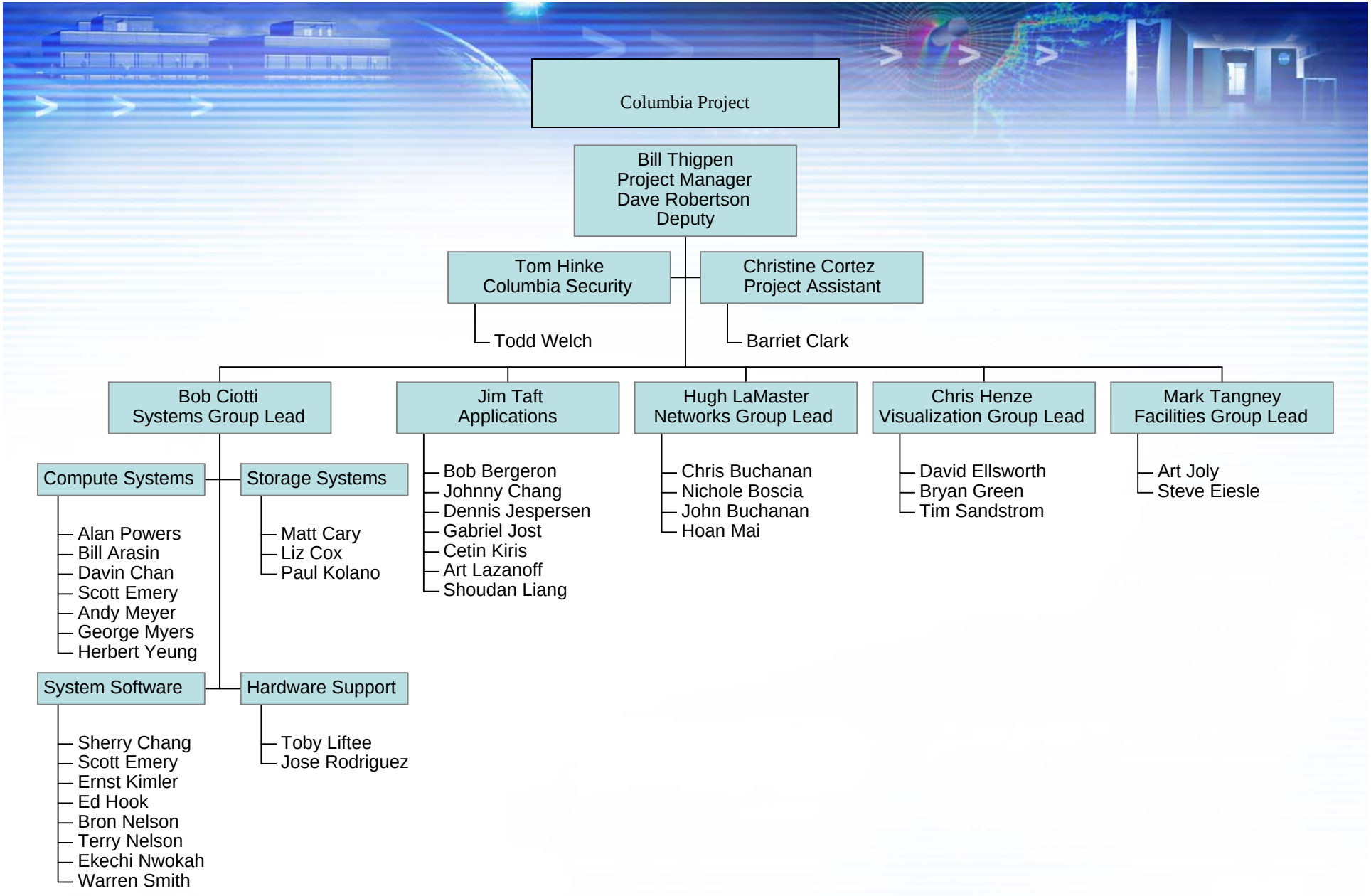
Project Columbia - Execution

The Impossible Tasks

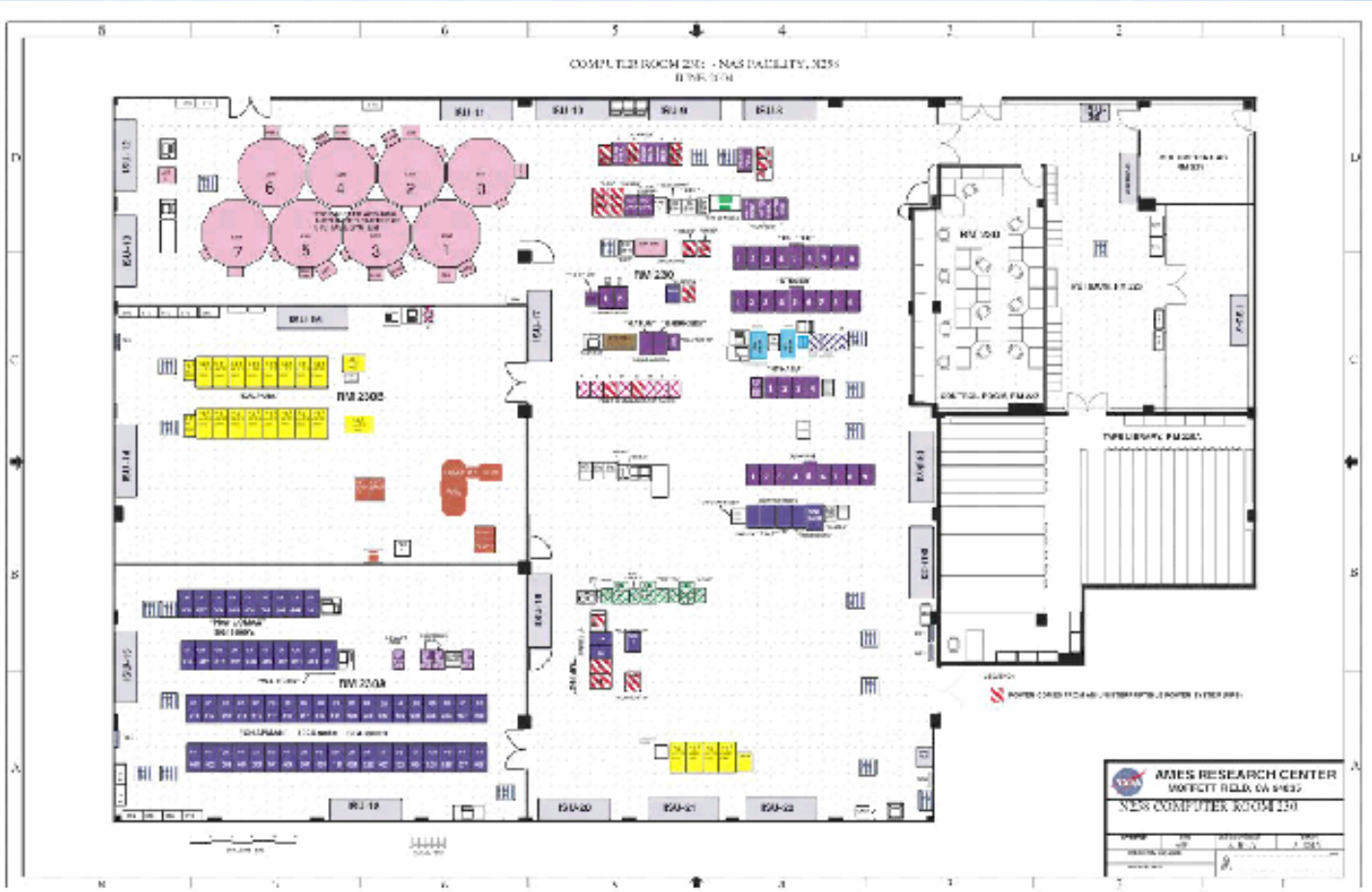
- Beginning in May 2004, obtain all necessary approvals and procure system in 1 month
- Make major upgrades to facility power and cooling
- Build and deliver Columbia in less then 4.5 months
- Eliminate negative slack on over 800 components
- To fit floor space, 8 nodes must use new double-density technology
- Continue and increase NASA mission science and engineering supercomputing during system build
- Beat benchmark speed of Earth Simulator system
- Build and utilize 1st shared-memory 2048 supercomputer in two weeks

Contributing Factors

- Teamwork
  - NASA Intel
  - NASA Intel SGI
  - NASA Intel SGI Voltaire
- Assume problems
  - Results focus
  - Eye on shared goal
  - Reduce waste
- Focus on Goals
  - Linpack options
  - Inter-company coordination


Columbia Project
Bill Thigpen
Project Manager
Dave Robertson
Deputy
Tom Hinke
Columbia Security
Todd Welch
Christine Cortez
Project Assistant
Barriet Clark
Bob Ciotti
Systems Group Lead
Jim Taft
Applications
Hugh LaMaster
Networks Group Lead
Chris Henze
Visualization Group Lead
Mark Tangney
Facilities Group Lead
Compute Systems
Alan Powers
Bill Arasin
Davin Chan
Scott Emery
Andy Meyer
George Myers
Herbert Yeung
Storage Systems
Matt Cary
Liz Cox
Paul Kolano
Bob Bergeron
Johnny Chang
Dennis Jespersen
Gabriel Jost
Cetin Kiris
Art Lazanoff
Shoudan Liang
Chris Buchanan
Nichole Boscia
John Buchanan
Hoan Mai
David Ellsworth
Bryan Green
Tim Sandstrom
Art Joly
Steve Eiesle
System Software
Sherry Chang
Scott Emery
Ernst Kimler
Ed Hook
Bron Nelson
Terry Nelson
Ekechi Nwokah
Warren Smith
Hardware Support
Toby Liftee
Jose Rodriguez

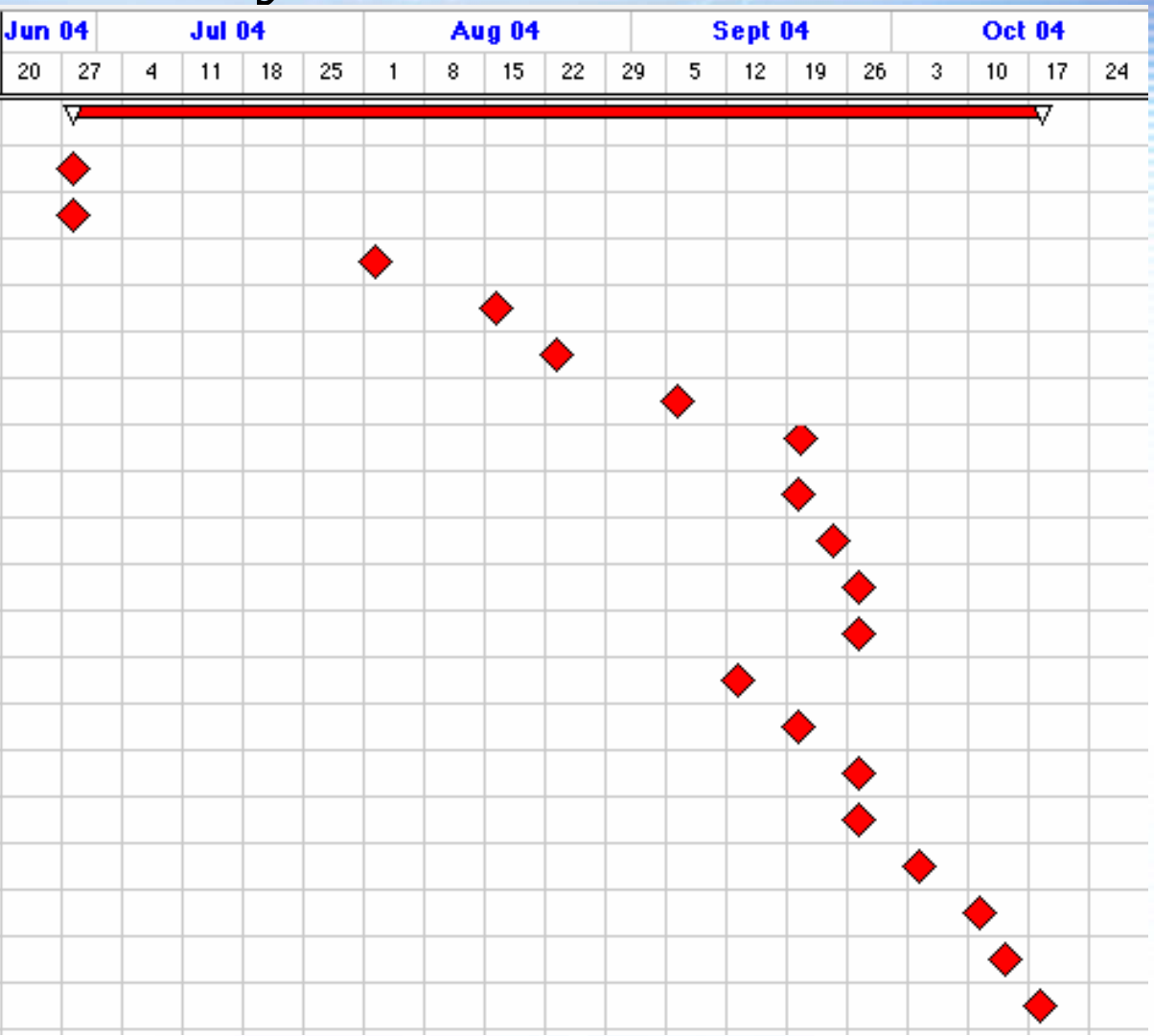
The Road There



June - 28 - 2004

Delivery Schedule

Activity Name	Start Date	Finish Date
System Arrival	6/28/04	10/18/04
3700-2	6/28/04	
3700-3	6/28/04	
3700-4	8/2/04	
3700-5	8/16/04	
3700-6	8/23/04	
3700-7	9/6/04	
3700-8	9/20/04	
3700-9	9/20/04	
3700-10	9/24/04	
3700-11	9/27/04	
3700-12	9/27/04	
3700BX2-1	9/13/04	
3700BX2-2	9/20/04	
3700BX2-3	9/27/04	
3700BX2-4	9/27/04	
3700BX2-5	10/4/04	
3700BX2-6	10/11/04	
3700BX2-7	10/14/04	
3700BX2-8	10/18/04	

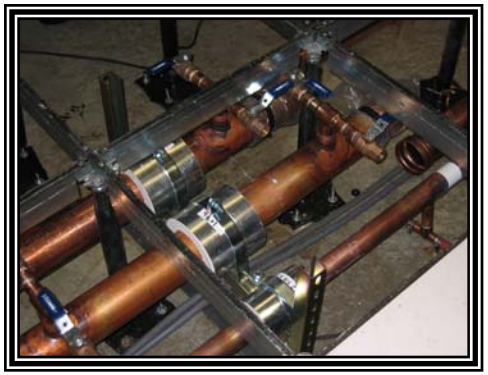


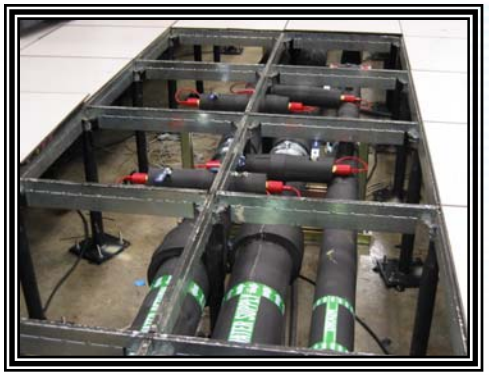
Power

- Successfully deployed cross-bar circuitry to allow for smoother power transitions in the future.
- 20 125kw PDU's received and installed
- New 800 amp distribution panels installed
- Power whips were installed prior to the arrival of system requiring the power

Cooling

- Floor Tiles received and installed as required
- Site visit to LLNL and Intel visit conducted
- Plumbing complete
- Heating problem contingency plans developed
- Water pipe burst survived
- Evaluating ceiling tile upgrades





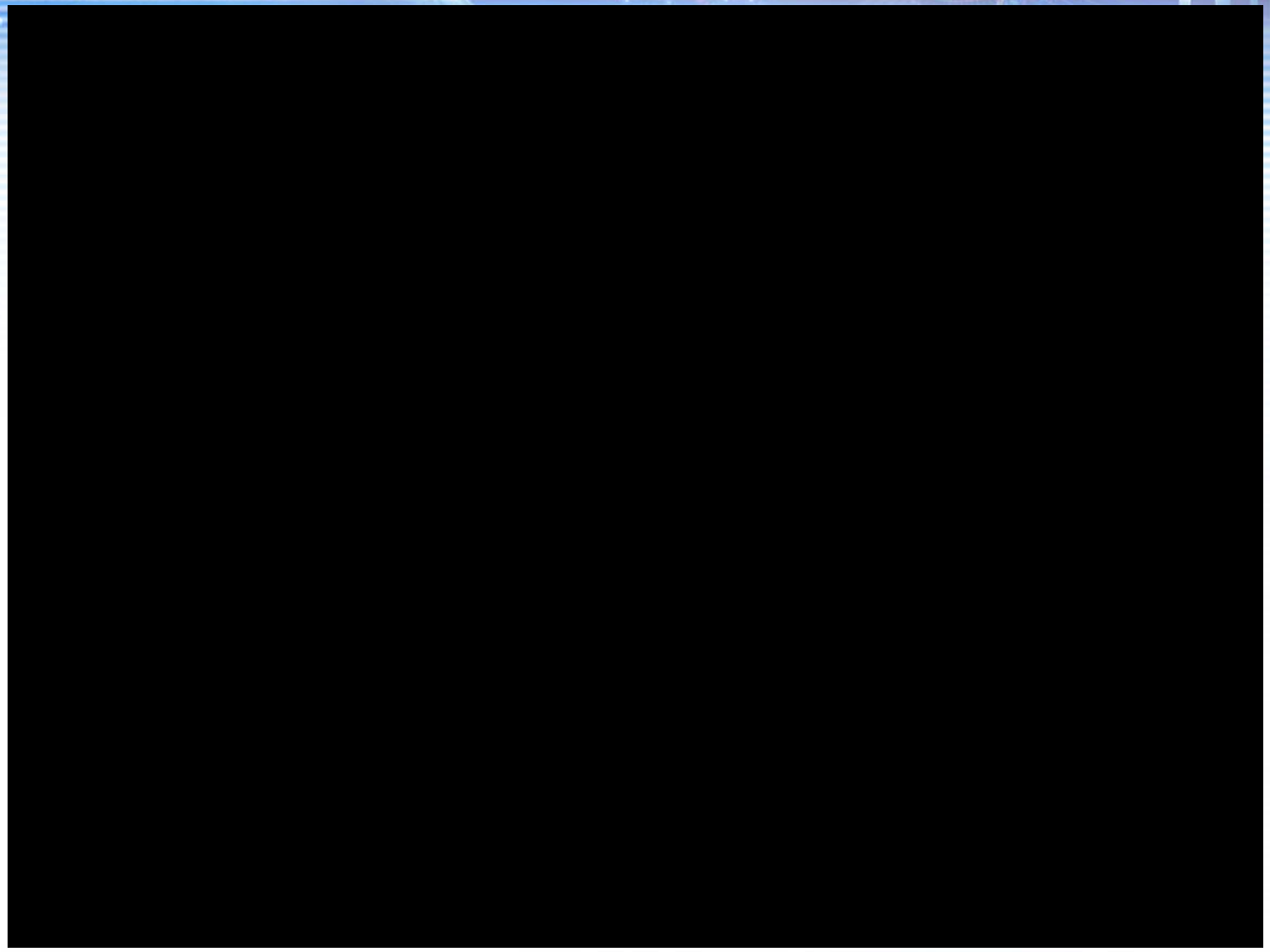
System Installation Highlights

- Each 512-processor node installed the day it arrived
- Hardware diagnostics begun that night
- Systems were often handed off from SGI to NAS systems personnel the next day
- OS and NAS specific hardware modifications added with disk in hours
- Initial connections were made with temporary cables
- Final configuration utilized cable infrastructure through under-floor or rack-mounted patch panels

Current Achievements

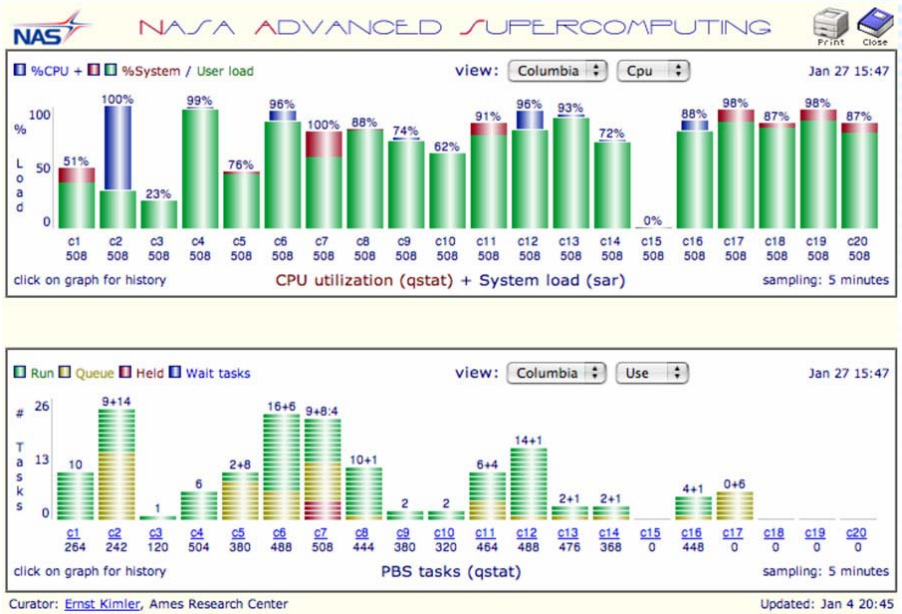
- The total system is in place
- Decommissioning and movement of the previous systems complete
- All facility modifications have been completed
- Disk farm in place
- Significant advancements on key applications in all NASA Mission areas
- Feedback from multiple users describes capabilities that were previously impossible
- Linpack runs were made with only two day interruption to RTF users and less then 1 week for all users
- 24-hour RTF rapid response was successfully completed
- 2048-processor system complete

We Are UP !!!



Columbia Is Operational

- 61TF Compute Capability
- A 64P node is the interactive front end
- The configuration is changeable to meet Agency requirements, current configuration:
  - 2 nodes each for the four mission directorates
  - 6 nodes for mission specified targeted applications
  - 4 nodes for 2048 scaling work
  - 2 nodes for systems development
- The chart depicts a normal load on the Columbia system providing the resource to make unprecedented strides in science and engineering



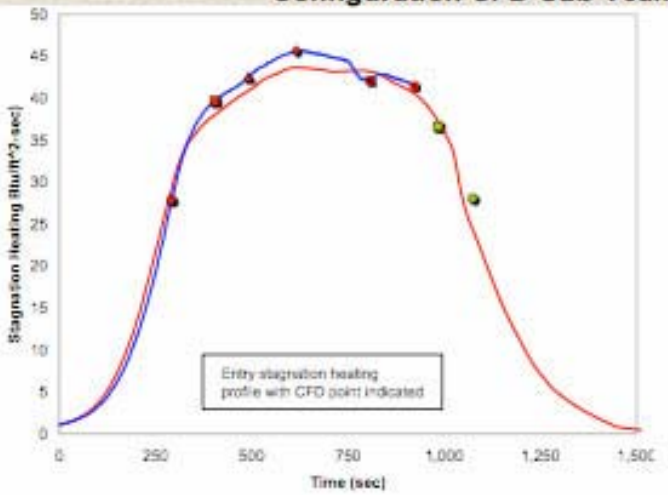
Rapid Aerothermal Analysis Demonstration

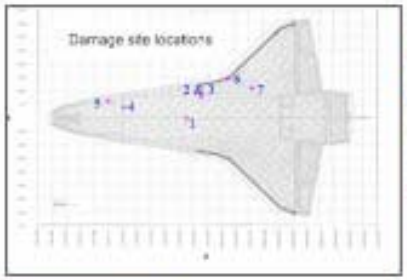
NASA RTF Computational Aerothermal Analysis Group: LaRC, ARC and JSC

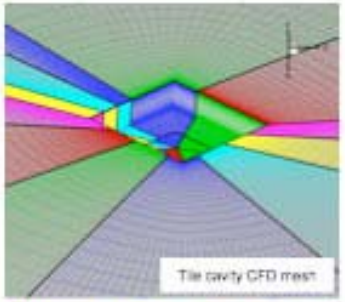
RTF Aerothermal Team ***Configuration CFD Sub-Team***

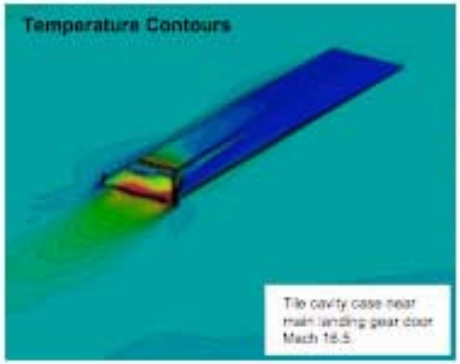
- **Objectives:**
  - Demonstrate a new rapid aerothermal CFD analysis capability
  - The new capability shall permit near real-time analysis of observed Orbiter damage during flight
  - The capability would provide an alternate high-fidelity evaluation of local heating bump factors calculated from engineering codes
  - Required capability: 10 damage sites at 10 trajectory points assessed in under 24 hours

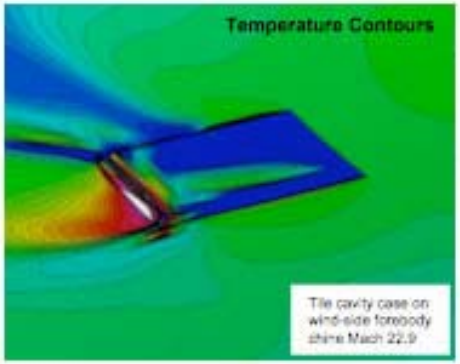
- **Results:** ***Success!!!***
  - 8 (512 processor) nodes of the *Columbia* supercomputer were used for 24 Hours
  - 12 different computational meshes were generated for 10 different damage/repair sites
  - 10 trajectory points calculated for each damage/repair site
  - Greater than 100 high fidelity Navier-Stokes calculations were performed
  - Some unsteadiness observed in some tile cavity CFD solutions
  - Some pre- and post processing improvements needed
  - Additional personnel would have been helpful

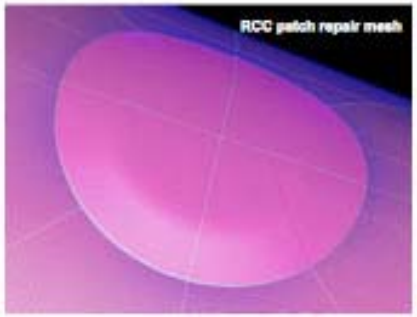


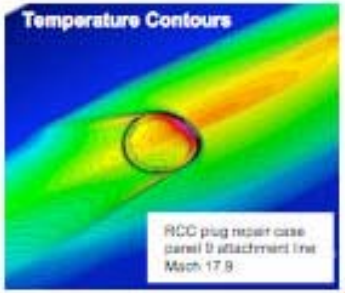






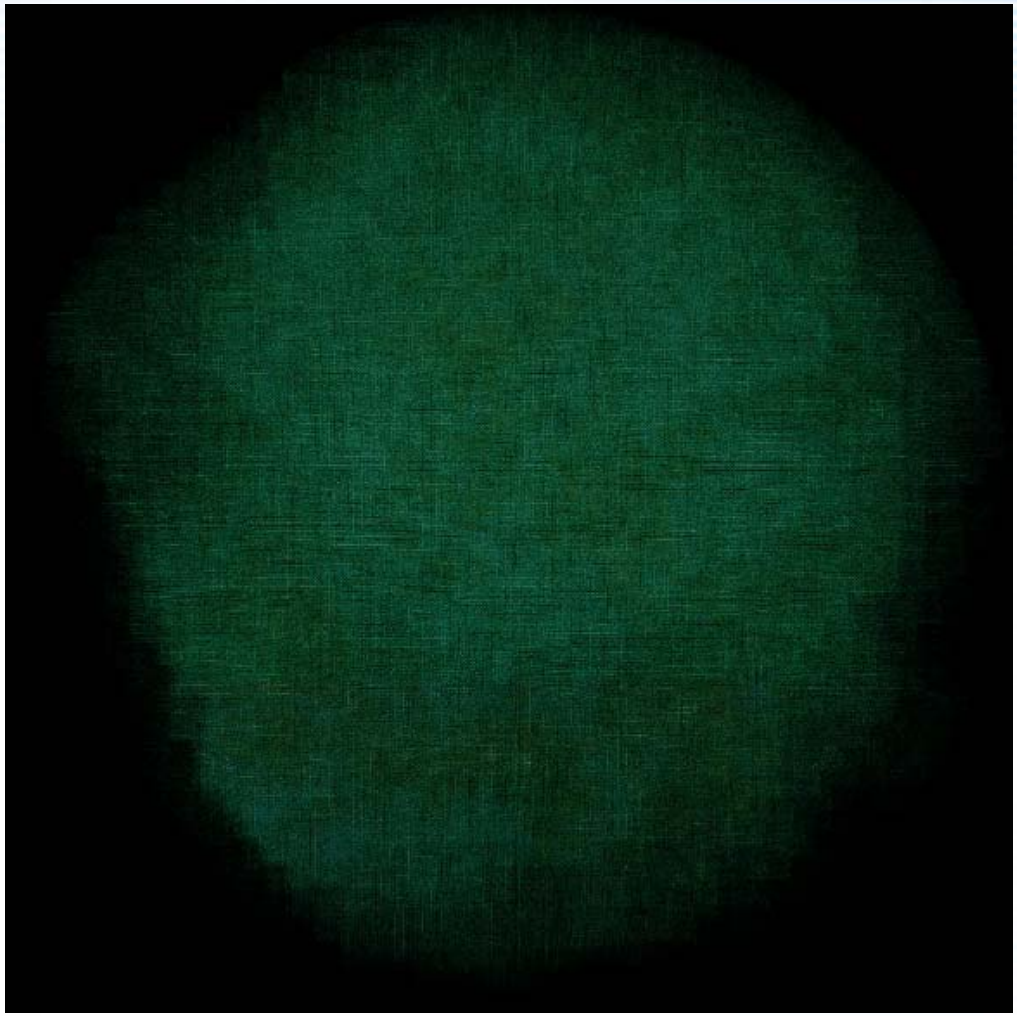




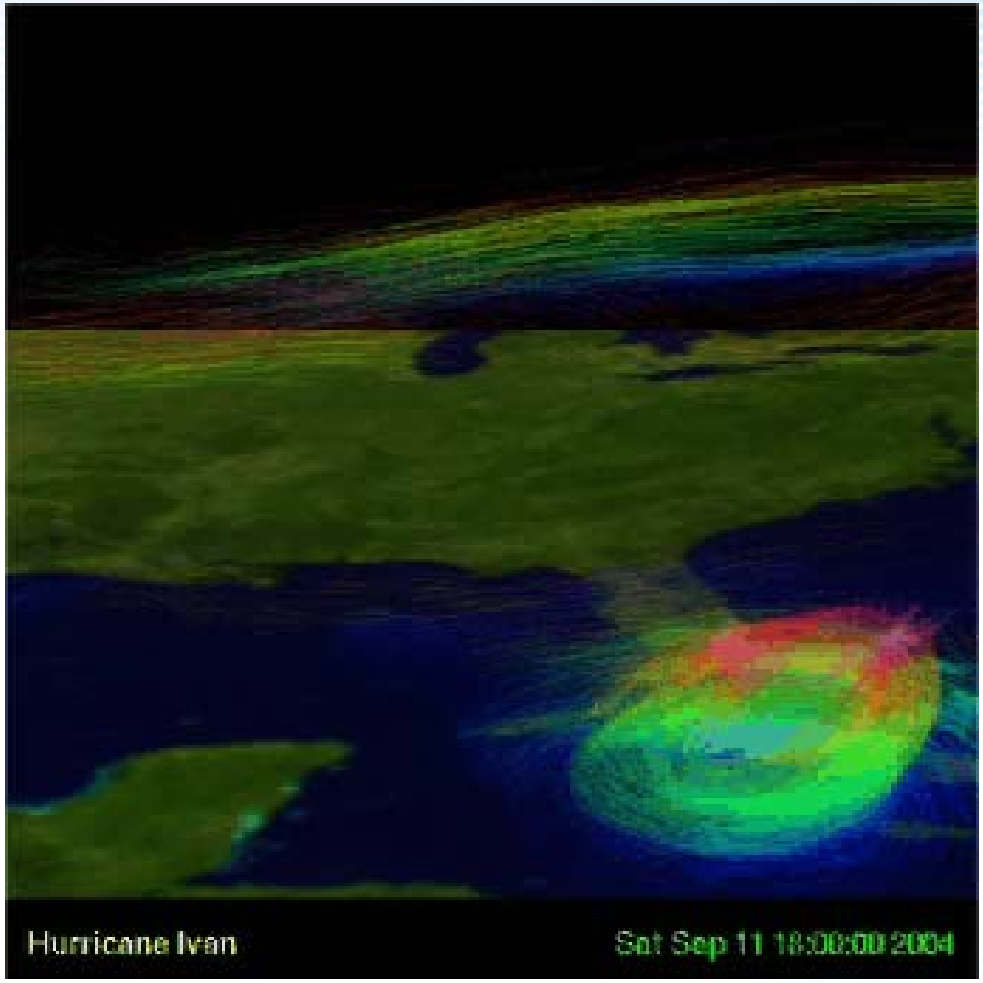


James Reuther, Peter Gnoffo, Dinesh Prabhu, Steve Alter, Ryan McDaniel, Maria Pulsonetti, Grant Palmer, David Saunders and Ben Kirk

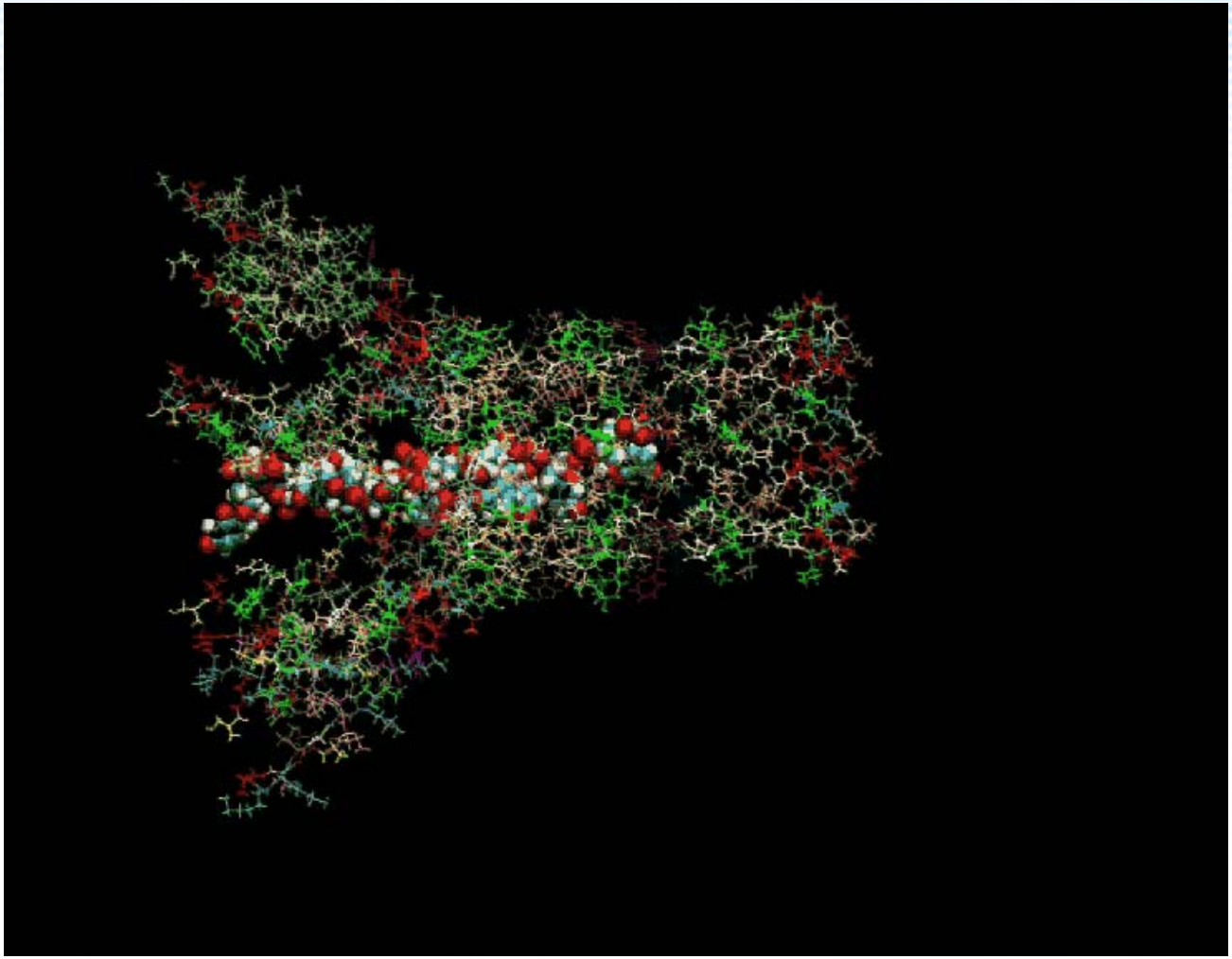
Columbia Animations



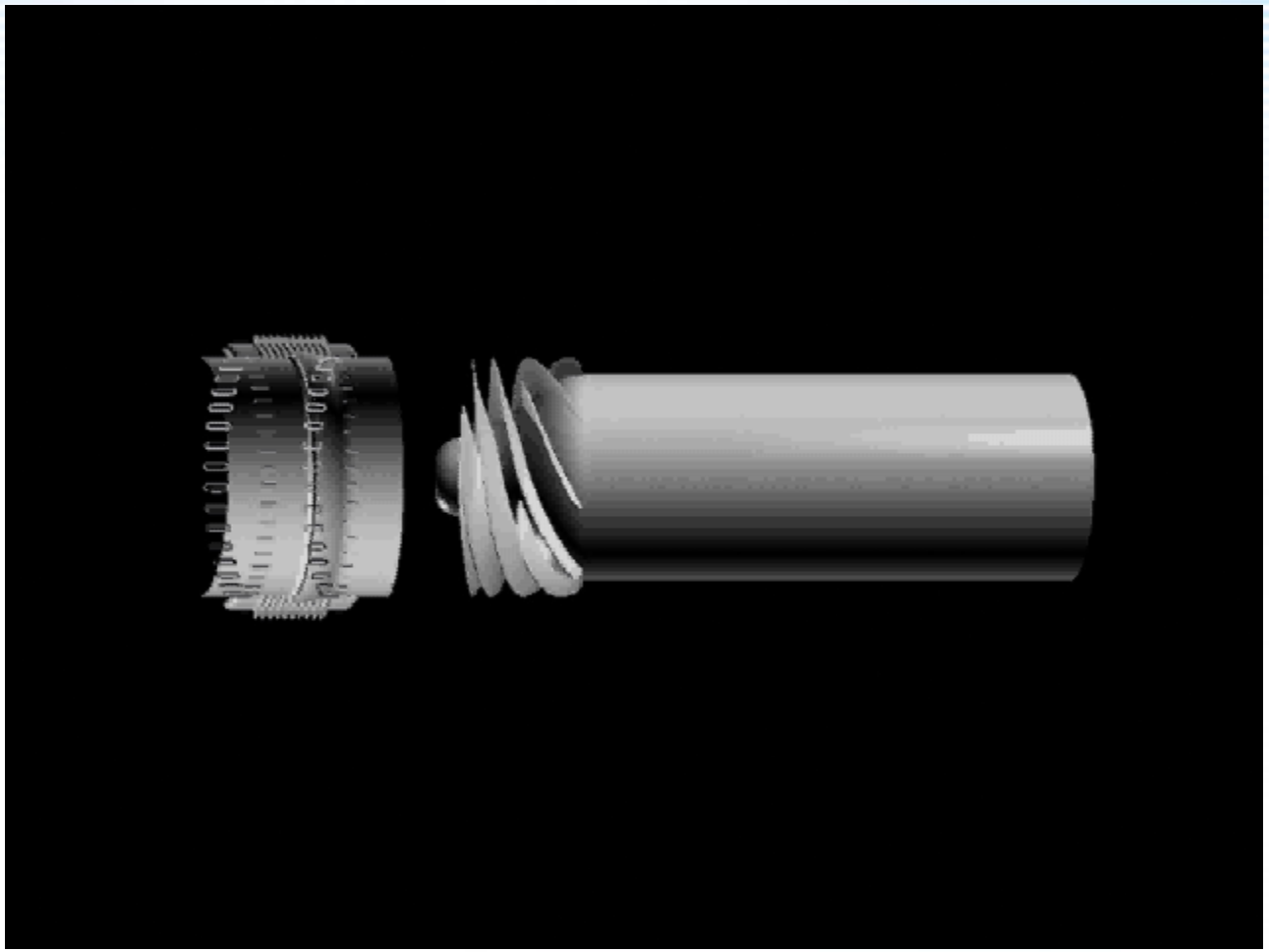
Columbia Animations



Columbia Animations



Columbia Animations



Summary – Columbia Supercomputer

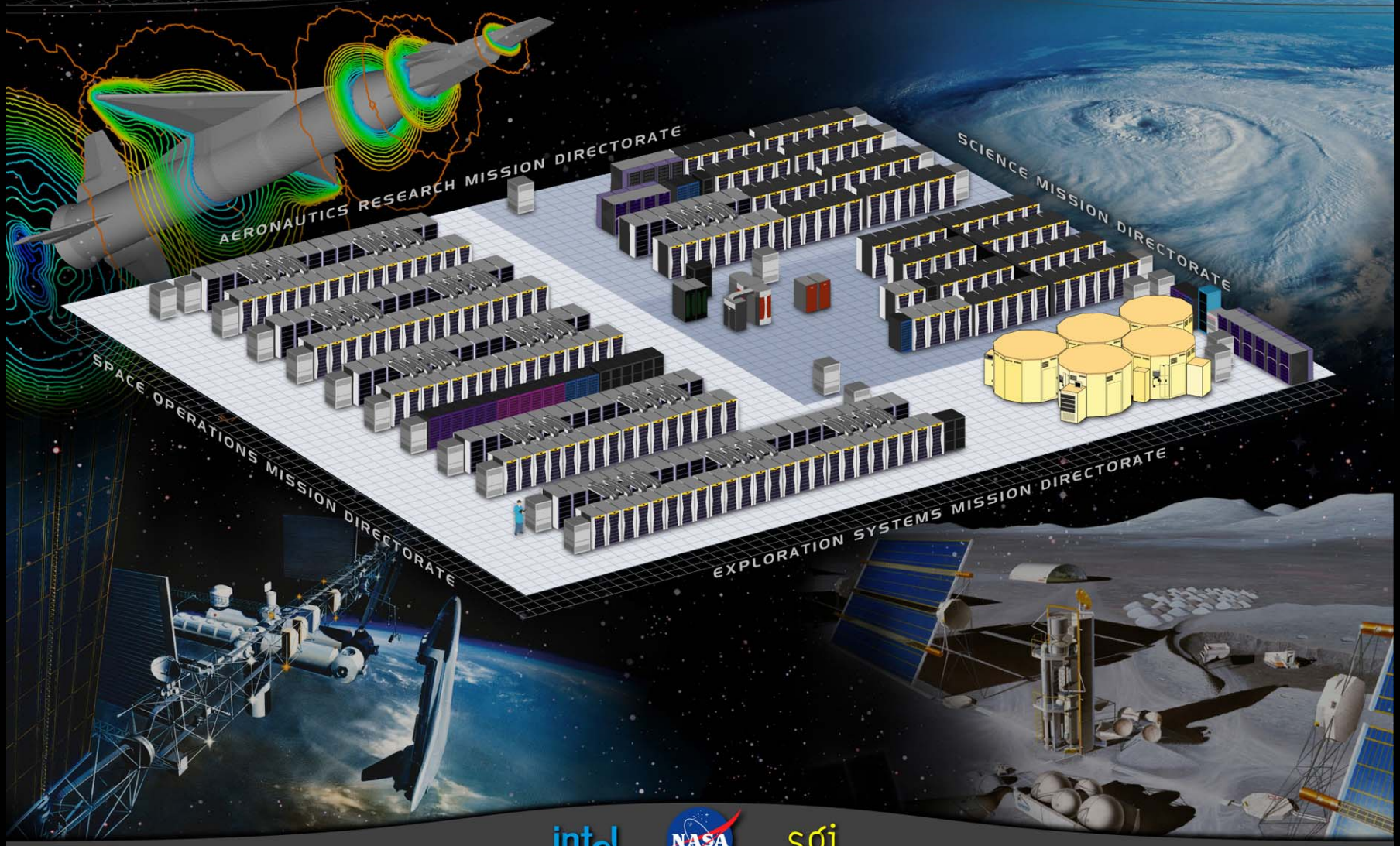
- World's Fastest Operational Supercomputer
- System Deployed in 109 Days (from delivery of first 2 512p nodes to full operations)
- Columbia is delivering unparalleled computational capability daily enabling new scientific discoveries
- Columbia stands ready to provide essential computing capabilities to meet rapid response requirements for Agency and National incidents
- 88% efficiency tops the scalar systems on the top500 list
- 2048 shared-memory environment has successfully supported initial application scaling

Summary – Project Columbia

- Fortune favors the bold, but dedication and flexibility are more reliable.
- Good partners build on each other's enthusiasm and performance.
- For time-critical work, true experts can save your bacon.
- It's never too late to anticipate – keep planning while executing.
- Communicate frequently but efficiently – tune your tools and techniques, target your talking.

PROJECT COLUMBIA

SUPPORTING NASA'S MISSION DIRECTORATES



intel. NASA sgi